APPENDIX B

WITNESS	TRANSCRIPT PINCITE	EXCERPT
Amanda Kallet	95:3-21[1]	Q. Do you know who "we" refers to here, "We most want to acquire"? A. "We" would be the LLM technical team, potentially -- I'm guessing -- yeah, the people who wrote it. Q. Okay. And it says "through either licensing or getting publicly available data approved for use." What does that mean? MS. DUNNING: Object to the form. Lacks foundation, calls for speculation. A. So before data can be used, it has to be approved internally. BY MS. GEMAN: Q. By whom? A. Legal mostly. Q. Who else? A. I'm not sure because I don't go through those processes. So beyond legal, I'm not sure.
Eleonora Presani	210:9-211:7[2]	Q. Who gave approvals for the use of data? MS. DUNNING: Object to the form. A. It was not just one person. The reason -- so what I was referring here is to go through the process of requesting approval, you have to go -- you have an internal review process and there's people assigned to it. Q. Sitting here today, do you recall some of those people who you had to request approvals from? MS. DUNNING: Object to the form. A. Well, I know that at least one of the people is within our legal department. Q. Is there anyone who's not in the legal department that you remember? A. There's a number of approvals for each of these, like, requests that go from communication to legal, but I think the last word is either with legal or -- yeah, I don't know. Legal, I would say.

---

[1] *See* Pritt Decl. Ex. K.
[2] *See* Pritt Decl. Ex. L.

1

WITNESS	TRANSCRIPT PINCITE	EXCERPT
Sy Choudhury	101:12-102:9[3]	Q. And you must have also considered whether Meta had the right to use those copyrighted works; correct?
		ATTORNEY HARTNETT: Object to form.
		THE WITNESS: We were working on engagements, external engagements, to determine what content was out there. It's our legal team that focuses on how to use the content. I'm not a lawyer.
		BY ATTORNEY PRITT:
		Q. So your legal team determines what content goes into the training data for use in Meta's large language models?
		ATTORNEY HARTNETT: Objection to form.
		THE WITNESS: I think our product and engineering work with a large cross-functional team, including legal; and so I'm not privy to those discussions or decisions. I just know that the -- the parts that are about the -- the third-party engagements and deals that we need to explore.
Sean Bell	70:7-19[4]	Q. And at the time that you joined the team, was Anna's Archive also an approved dataset for training Llama 4?
		A. I --
		ATTORNEY HARTNETT: Object to the form.
		THE WITNESS: I don't think we had an explicit decision one way or the other. And so Anna's Archive was data -- was a superset of LibGen. And so the decision that I had from legal was that LibGen was an approved data source, and that -- and so it didn't -- at the time I joined, I don't recall it contemplating one way or another about Anna's Archive.
Chaya Nayak	93:16-21[5]	Q. And who -- who works in the data approval process? Who -- who -- who are some of the folks that you were working with that were part of the approval process?
		A. Our legal team, our privacy team, our engineers, and our product team.

---

[3] *See* Pritt Decl. Ex. M.
[4] *See* Pritt Decl. Ex. N.
[5] *See* Pritt Decl. Ex. O.

WITNESS	TRANSCRIPT PINCITE	EXCERPT
Alex Boesenberg	383:5-384:12[6]	Q. Okay. So on April 7, 2023, you testified that Mr. Choudhury issued a directive to cease negotiation efforts to license – cease negotiation efforts with third parties to license the materials for training AI data; right? A. Yes. Q. Okay. Do you recall why that directive was issued? A. Yes. MR. WEINSTEIN: Caution the witness just -- go ahead. BY MR. YOUNG: Q. Okay. Why was that directive issued? MR. WEINSTEIN: Caution the witness not to reveal -- THE WITNESS: Subject to -- MR. WEINSTEIN: Just a minute. Caution the witness not to reveal the content of any -- THE WITNESS: Yeah, I mean -- MR. WEINSTEIN: Let me finish. I know it's late in the day. Caution the witness not to reveal the content of any communication with counsel. Otherwise, you can answer. THE WITNESS: Yeah, I mean, I think that there was a discussion between Sy and legal leadership, and there was a conclusion that it didn't make sense for us to further pursue.

---

[6] *See* Pritt Decl. Ex. P.

WITNESS	TRANSCRIPT PINCITE	EXCERPT
Melanie Kambadur	117:13-118:14[7]	Q. Now, do you remember if Meta had the Libgen dataset before it was processed and including in the Llama 3 models?
		MR. WEINSTEIN: Object to form.
		A. Of course we had to have the dataset before including it in the model, yes.
		BY MR. YOUNG:
		Q. So that means that there was a decision made to incorporate Libgen, process Libgen as pre-training date for Llama 3, correct?
		A. Yes.
		Q. Do you know who made that decision?
		A. I don't think there was a singular -- there was not a singular person.
		Q. But a decision was made, correct?
		A. To include Libgen into Llama 3? Yes.
		Q. Do you remember why that decision was made to include Libgen into Llama 3?
		A. As with all of our datasets, we had run some small-scale experiments that indicated it would positively improve our benchmarks on a larger model. And we had gone through our legal and privacy review procedures.

---

[7] *See* Pritt Decl. Ex. Q.

WITNESS	TRANSCRIPT PINCITE	EXCERPT
Joelle Pineau	55:21-57:12[8]	Q. Would it be fair to say after September or October 2022 when those individuals entered your management chain, that you would have the ability to either approve or disapprove of datasets they were considering using for training AI models?
		MS. HARTNETT: Objection to form.
		A. The approval process for datasets is a little more complicated than that. It's a shared responsibility between research and privacy team and our legal team.
		BY MR. YOUNG
		Q. But you would have input on whether or not a certain dataset would be approved or disapproved?
		MS. HARTNETT: Objection to form.
		A. So usually the research director has a point of view on the content of the research project; so whoever is the first line research director is the approver on the research plan. And the datasets are specified as part of that plan. And so we have a, you know essentially a voice into the approval process, but all the deciders need to align so it's not a singular approval.
		BY MR. YOUNG
		Q. Have there been any examples of any datasets that you have expressed concerns about using for AI training during your time at Meta?
		MS. HARTNETT: Objection to form.
		A. I have definitely asked questions about a number of datasets over time and discussed this with my legal and privacy partners.

---

[8] *See* Pritt Decl. Ex. R.